AD-A085 885 YALE UNIV NEW HAVEN CONN DEPT OF STATISTICS F/G 12/1
TESTS FOR JOINT NORMALITY IN TIME SERIES.(U)
MAY 80 F J ANSCOMBE, H H CHANG N00014-75-C-0563
UNCLASSIFIED TR-47 NL

1 of 1

AD A
085885

END
DATE
FILMED
8-80
DTIC

YALE UNIVERSITY

Department of Statistics

Technical Report No. 47

May, 1980

F. J. Anscombe and H. H. Chang

"Tests for joint normality in time series"

Research supported by the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Contract number: N00014-75-C-0563

Task number: NR 042-242

TESTS FOR JOINT NORMALITY IN TIME SERIES

1. Motivation

The well-known methods for analysis of time series, whether in the time domain or in the frequency domain -- for fitting parametric structures, for regression, for forecasting -- all involve second-moment statistics. If all variables are jointly normally distributed in stationary sequences, simple first and second moments contain all the information. If not, there is the possibility that some of the needed information is not contained in the statistics used. When a random sequence is other than stationary and jointly normal, it may sometimes equally well be described and thought of as stationary but not jointly normal (which is the terminology used here) or as nonstationary.

The topic is most easily illustrated in the context of simple regression. Suppose that we desire to be able to estimate the unobserved value of y when x is observed, that x and y are random variables with a joint distribution, and that we have a large sample of independent (x, y) observations with which to estimate the relation between y and x. If the joint distribution of x and y is normal, first and second moments of the sample are sufficient statistics, the regression curve of y on x (that is, the conditional expectation of y, given x) is linear and well estimated, for many purposes, by the usual least-squares regression line. In particular, the ordinate of the fitted line is the best estimate of y, given x, in the absence of prior information about the parameters (that is, for a flat prior) and for any loss function that is not constant and is a nondecreasing function of the magnitude of the error.

When the joint distribution for x and y is not normal, the usual regression line may be less satisfactory. Consider three examples.

(i) If the regression curve of y on x is not linear, fitting a linear regression relation to some data may give a poor way of estimating y for given x. Let

$$y = \mu + \beta|x| + \varepsilon ,$$

where μ and β are constants, x is distributed $N(0, 1)$ and ε is distributed independently $N(0, \sigma^2)$, and $\sigma^2 < \beta^2$. The usual regression line is useless and suggests a much larger residual variance than the correct σ^2. Here the marginal distribution for y is not normal.

(ii) If x and y are marginally normal and are jointly distributed with constant crossproduct probability ratio different from 1 (Plackett, 1965), the regression curve of y on x is monotone but not linear. This is a less dramatic example of the same effect as at (i); the usual regression line is not useless, but it is not the best predictor.

(iii) Even if the regression curve of y on x is linear, the usual regression line may give a poor estimate of y, given x, if the loss function is sufficiently different from squared error. Let x be distributed $N(0, 1)$ and, given x, let y have probability $\frac{1}{2}$ of being equal to x, and probability $\frac{1}{2}$ of being an independent $N(0, 1)$ variable. The regression curve of y on x is then $y = \frac{1}{2}x$. Let the loss be 0 if $|\text{error}| \simeq \delta$, otherwise 1, where δ is small. Then for given x, if say $x > 0$, a good estimate of y is $\max[0, x - \delta]$. Here, as for (ii), the marginal distributions for x and y are normal.

These considerations do not seem very sinister in regard to simple linear regression, because a scatterplot of the given (x, y) observations would most likely reveal any such effects. No special machinery seems to be called for. Similar considerations apply to multiple regression on several explanatory variables. Obtaining a comprehensive understanding of how the variables are related through examining scatterplots is less easy, though still possible. Various tests can also be calculated from the residuals.

Perceiving nonnormality in the joint distribution of a single time series, or of several related time series, is difficult from such graphical displays as are commonly made in treating time series. A correlogram or spectrum (or cross-correlogram or cross-spectrum) does not help much. We here propose an adaptation to stationary time series of Mardia's test for kurtosis in a multivariate distribution. In this preliminary report, suitable test statistics are proposed, some information is given concerning their distribution under the null hypothesis, with a suggested computer program for making the tests, and there is a brief consideration of power. Further study of these matters, and examples of application, will be presented later.

2. Tests for kurtosis in stationary time series

A given time series $\{x_t\}$, for some consecutive integer values for t , is supposed to be realized from a stationary sequence of random variables $\{\xi_t\}$, and we wish to test the hypothesis that the random variables are jointly normally distributed.

Univariate normality of the marginal distribution for ξ_t could be tested by making a histogram of the aggregate of given values $\{x_t\}$ or by calculating (for example) a kurtosis statistic,

$$b_2 = N(\Sigma_t(x_t - \bar{x})^4)/(\Sigma_t(x_t - \bar{x})^2)^2 ,$$

where N is the number of t-values and $N\bar{x} = \Sigma_t x_t$. To determine a significance level for b_2 , proper account would have to be taken of the correlation structure of the sequence $\{\xi_t\}$.

Univariate marginal normality of a stationary random sequence does not imply joint normality, and the latter is what we are interested in here. Mardia (1970) has considered testing joint normality, given n independent observations of a p-variate distribution. His procedure involves linearly transforming the p-variate distribution so that it becomes spherical, and then he considers the n distances of the observations from their center of gravity and constructs a kurtosis statistic by comparing the sum of the fourth powers of the distances with the squared sum of the squared distances. An analogous way to treat a stationary time series would be to express the series in terms of independent identically distributed "innovations" and then calculate kurtosis statistics either from single innovations, or from pairs of consecutive innovations, or from triples of consecutive innovations, etc.

Our suggested procedure is, first of all, to try to represent the sequence in finite autoregressive form, say

$$(\xi_t - \mu) - \alpha_1(\xi_{t-1} - \mu) - \cdots - \alpha_p(\xi_{t-p} - \mu) = \varepsilon_t \quad (t = 0, \pm 1, \pm 2, \ldots) , \tag{1}$$

where $\mu, \alpha_1, \ldots, \alpha_p$ are constants and $\{\varepsilon_t\}$ are independent identically distributed "error" random variables having a normal distribution with zero mean. For a given positive integer p , such a finite autoregressive structure can be estimated by performing ordinary linear regression of $\{x_t\}$ on $\{x_{t-1}\}, \{x_{t-2}\}, \ldots, \{x_{t-p}\}$, where we may say that $1 \leq t \leq n$ if the whole given series is of length $n + p$, that is, the given series is $\{x_t\}$ for $1 - p \le t \le n$. The residuals $\{u_t\}$ are the "innovations", estimating the errors $\{\epsilon_t\}$:

$$u_t = x_t - a_0 - a_1 x_{t-1} - \cdots - a_p x_{t-p} \quad (1 \le t \le n) , \tag{2}$$

where $a_0, a_1, \ldots, a_p$ are the regression coefficients. The innovations depend on the choice of p. A possible method of choosing p is to calculate the empirical discrete spectrum of the given series (prewhitened and tapered), smooth it with a suitable moving average to form (after adjusting for the prewhitening) an estimated spectral density, and then try to approximate the reciprocal of the spectral density by a low-order polynomial in the cosine of the angular frequency; p is taken to be the degree of the polynomial. We shall suppose that n is much larger than p. Formation of innovations has been recently discussed by Kleiner, Martin and Thomson (1979), for a different purpose.

For a given vector of innovations $\{u_t\}$, a kurtosis statistic can be defined from single innovations:

$$b_{21} = n\Big(\sum_{t=1}^{n} u_t^4\Big)\Big/\Big(\sum_{t=1}^{n} u_t^2\Big)^2 . \tag{3}$$

A kurtosis statistic defined from pairs of consecutive innovations is

$$b_{22} = n\Big(\sum_{t=1}^{n-1} (u_t^2 + u_{t+1}^2)^2\Big)\Big/\Big(\sum_{t=1}^{n} u_t^2\Big)^2 , \tag{4}$$

and one based on triples of consecutive innovations is

$$b_{23} = n\Big(\sum_{t=1}^{n-2} (u_t^2 + u_{t+1}^2 + u_{t+2}^2)^2\Big)\Big/\Big(\sum_{t=1}^{n} u_t^2\Big)^2 , \tag{5}$$

and so on.

If indeed the sequence $\{\xi_t\}$ is correctly described by an expression of the form (1), for some finite p, the left side of (1) is a sequence of independent identically distributed normal variables. Then if the correct value for p is used, the innovations $\{u_t\}$ will presumably seem to be realized from nearly independent identically distributed normal variables and the kurtosis statistics $b_{21}, b_{22}, \ldots$ should behave accordingly. In particular, if n is large, b_{21} is expected to be near to $E(\epsilon_t^4)/(E\epsilon_t^2)^2$, which is equal to 3, b_{22} is expected to be near to $E((\varepsilon_t^2 + \varepsilon_{t+1}^2)^2)/(E\,\varepsilon_t^2)^2$, which is equal to 8, b_{23} is expected to be near to 15, etc.

If the sequence $\{\xi_t\}$ is jointly normal but, for some p, cannot be represented in the form (1), either because a larger value for p would be needed or because the sequence does not have a finite autoregressive expression, and if for the chosen value of p the parameters $\mu, \alpha_1, \ldots, \alpha_p$ are chosen to minimize the variance of the left side of (1), then the left side constitutes a stationary sequence of normal variables that are not independent. The innovations calculated for that p will presumably also seem to be realized from correlated normal variables. Correlation in the innovations may be expected to have less effect on b_{21} than on $b_{22}, b_{23}, \ldots$

3. Distributions under the null hypothesis

To approximate the distributions of the statistics b_{21}, b_{22}, etc. under the null hypothesis of stationarity and joint normality, it is natural to consider moments. It has been found that the distribution of the ordinary kurtosis statistic, Pearson's b_2 or Fisher's g_2, in samples from a normal population, is fairly well approximated by a linear function of the reciprocal of a χ^2 variable, having a distribution of Pearson's Type V, fitted to the first three moments (Anscombe and Glynn, 1975). Accordingly we seek to determine the first three moments of the distributions of b_{21}, b_{22}, ..., in order to be able to make the Type V approximation.

First suppose that $\{u_t\}$ in the definitions (3), (4) and (5) are not as specified at (2) but instead are independent $N(0, 1)$ variables. Since the ratio on the right side of each definition is then independent of its denominator, relations such as this hold:

$$E(b_{21}{}^r)\, E(\Sigma_t\, u_t{}^2)^{2r} = n^r\, E(\Sigma_t\, u_t{}^4)^r \quad (r = 1, 2, \ldots) .$$

The following results may be deduced.

For $n \geq 1$,

$$E(b_{21}) = \frac{3n}{n+2} ,$$

$$\operatorname{var}(b_{21}) = \frac{24\, n^2(n-1)}{(n+2)^2(n+4)(n+6)} \sim \frac{24}{n+15} ,$$

$$E(b_{21} - Eb_{21})^3 = \frac{1728\ n^3(n-1)(n-2)}{(n+2)^3(n+4)(n+6)(n+8)(n+10)} \sim \frac{1728}{n(n+37)},$$

$$\gamma_1(b_{21}) \sim \frac{6\sqrt{6}}{\sqrt{n+29}} \doteq \frac{14.70}{\sqrt{n+29}} .$$

The asymptotic results after the $\sim$ sign are correct to a factor $1 + O(n^{-2})$ when n is large. The skewness measure γ_1 means the third central moment divided by the standard deviation cubed.

Similarly for $n \geq 2$,

$$E(b_{22}) = \frac{8(n-1)}{n+2} \sim \frac{8n}{n+3},$$

$$\text{var}(b_{22}) = \frac{16(n-2)(7n^2+2n+48)}{(n+2)^2(n+4)(n+6)} \sim \frac{112}{n+15\frac{5}{7}},$$

and for $n \geq 3$,

$$E(b_{22} - Eb_{22})^3 = \frac{256(65n^5 - 358n^4 + 996n^3 - 1928n^2 + 5152n - 7680)}{(n+2)^3(n+4)(n+6)(n+8)(n+10)}$$

$$\sim \frac{16640}{n(n+39\frac{33}{65})},$$

$$\gamma_1(b_{22}) \sim \frac{260}{7\sqrt{7}\sqrt{n+31\frac{397}{455}}} \doteq \frac{14.04}{\sqrt{n+31.9}} .$$

For $n > 3$,

$$E(b_{23}) = \frac{15(n-2)}{n+2} \sim \frac{15n}{n+4},$$

and for $n \geq 4$,

$$\text{var}(b_{23}) = \frac{8(37n^3 - 3n^2 + 296n - 2700)}{(n+2)^2(n+4)(n+6)} \sim \frac{296}{n+14\frac{3}{37}},$$

and for $n > 6$,

$$E(b_{23} - Eb_{23})^3 = \frac{192(365n^5 - 2709n^4 + 9414n^3 - 8516n^2 + 79800n - 540000)}{(n+2)^3(n+4)(n+6)(n+8)(n+10)}$$

$$\sim \frac{70080}{n(n+41\frac{154}{365})},$$

$$\gamma_1(b_{23}) \sim \frac{13.76}{\sqrt{n+40.6}} \text{ (approximately).}$$

Determination of these expressions has been partly computerized, and there is reason to hope that correctness has been achieved. The skewness measure (γ_1) of the distribution of each of the statistics b_{21}, b_{22}, b_{23} is nearly the same, when n is large. This suggests that the shapes of the distributions may be similar.

It is possible in principle to obtain exact expressions for moments of the statistics, supposing that $\{u_t\}$ in the definitions (3), (4), (5) are residuals from the fitting by least squares of a linear regression relation on given explanatory variables, the errors being independent $N(0, 1)$ variables (Anscombe, 1961). The exact expressions depend on the projection matrix $\underset{\sim}{Q}$ that transforms the errors into the residuals, but under mild conditions on the explanatory variables the asymptotic expressions quoted above, correct to a factor $1 + O(n^{-2})$, remain valid.

For example, if only a general mean is estimated, the $\{u_t\}$ are independent $N(0, 1)$ variables with their average subtracted. Results for b_{21} in this case, for $n \geq 2$, are due essentially to Fisher (1930):

$$E(b_{21}) = \frac{3(n-1)}{n+1} \sim \frac{3n}{n+2},$$

$$\operatorname{var}(b_{21}) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)} \sim \frac{24}{n+15},$$

$$E(b_{21} - Eb_{21})^3 = \frac{1728n(n-2)(n-3)(n^2-5n+2)}{(n+1)^3(n+3)(n+5)(n+7)(n+9)} \sim \frac{1728}{n(n+37)}.$$

The asymptotic results are the same as before. Now let $\{u_t\}$ be residuals from the fitting of a general mean and also a regression coefficient on an explanatory variable $\{z_t\}$. Let $\{z_t\}$ be scaled by subtracting the average and dividing by the square root of the sum of squares (assumed to be positive). Then $\Sigma_t z_t = 0$, $\Sigma_t z_t^2 = 1$. We require that uniformly for every t, $z_t = O(n^{-\frac{1}{2}})$; in particular, $\Sigma_t z_t^4 = O(n^{-1})$. This will happen in probability if $\{z_t\}$, before the scaling just mentioned, was a random sample from some stationary random sequence having positive variance and all moments finite. The condition prevents the regression coefficient on $\{z_t\}$ from being largely determined by a single reading. We find (for $n \geq 3$)

$$E(b_{21}) = \frac{3}{n-2}\left\{\frac{(n-1)(n-3)}{n} + \Sigma_t z_t^4\right\} \sim \frac{3n}{n+2}.$$

Similar exact expressions can be exhibited for other moments, though much clumsier.

In fact, our time-series innovations $\{u_t\}$ defined at (2) above are residuals from the fitting by least squares, not of a linear regression relation on given explanatory variables, but of a linear autoregression relation. We conjecture, but have not proved, that the same asymptotic results hold.

Below is given an APL program for making the b_{21}, b_{22} and b_{23} tests on a given vector of innovations. Equivalent normal deviates (E.N.D.) are calculated from the conjectured asymptotic moments, the Type V approximation and the Wilson-Hilferty approximation to the distribution of χ^2 .

```
     ∇ TSNT U;B;D;E;M;N;S
[1]    →2+3=⎕NC 'END'
[2]    →0,ρ⍞←'COPY END FROM 1234 ASP3'
[3]    →4+(∧/,0<D←(+/S←U×U)*2)∧(1≤N←+/ρU)∧1=ρρU
[4]    →0,ρ⍞←'NO GO.'
[5]    'B21  =  ',⍕B←(N×S+.×S)÷D
[6]    'APPROXIMATE MEAN  =  ',(⍕M←3÷1+2÷N),',  S.E.  =  ',(⍕E←(24÷N+15)*÷2),',
       GAMMA1  =  ',⍕G←14.7÷(N+29)*÷2
[7]    '   (E.N.D.  = ',(2⍕END),')'
[8]    'B22  =  ',⍕B←(N×+/((1↓S)+¯1↓S)*2)÷D
[9]    'APPROXIMATE MEAN  =  ',(⍕M←8÷1+3÷N),',  S.E.  =  ',(⍕E←(112÷N+15.7)*÷2),'
       ,  GAMMA1  =  ',⍕G←14.04÷(N+31.9)*÷2
[10]   '   (E.N.D.  = ',(2⍕END),')'
[11]   'B23  =  ',⍕B←(N×+/((2↓S)+(1↓¯1↓S)+¯2↓S)*2)÷D
[12]   'APPROXIMATE MEAN  =  ',(⍕M←15÷1+4÷N),',  S.E.  =  ',(⍕E←(296÷N+14.1)*÷2),
       ',  GAMMA1  =  ',⍕G←13.76÷(N+40.6)*÷2
[13]   '   (E.N.D.  = ',(2⍕END),')'
[14] ⍝  TIME SERIES NORMALITY TEST.  THE ARGUMENT IS A VECTOR OF INNOVATIONS.
     ∇


     ∇ X←END;A
[1]    A←6+4×A×A+4○A←2÷G
[2]    →3+0<X←1+(B-M)÷E÷(2÷A-4)*÷2
[3]    →0,ρ⍞←'E.N.D. NOT FOUND.',X←''
[4]    X←(1-(2÷9×A)+((1-2÷A)÷X)*÷3)÷(2÷9×A)*÷2
[5]  ⍝  INVOKED BY 'TSNT'.
     ∇
```

4. Power considerations

The tests are intended to be responsive to nonnormality in the joint distribution of the random sequence. They should preferably respond little to specification error in a jointly normal random process, that is, to choosing too low a value for the order p of the autoregressive structure fitted.

Suppose that p is chosen to be 1. For present purposes the mean of the sequence may be set equal to 0. Then, if p is correct, the null hypothesis is that $\{\xi_t\}$ is a jointly normal stationary Markov sequence:

Hypothesis A: $\xi_t = \rho\xi_{t-1} + \varepsilon_t$, where ρ is constant $(|\rho| < 1)$ and ε_t is distributed $N(0, 1 - \rho^2)$ independently of ξ_{t-1}, $\xi_{t-2}, \ldots$

An alternative hypothesis involving marginal normality but not joint normality is that $\{\xi_t\}$ is a stationary Markov "jump" sequence:

Hypothesis B: With probability ρ/a , $\xi_t = a\xi_{t-1} + \varepsilon_t$, when a and ρ are constant $(0 < \rho < a \leq 1)$ and ε_t is distributed $N(0, 1 - a^2)$ independently of $\xi_{t-1}, \xi_{t-2}, \ldots$; and with probability $1 - \rho/a$, $\xi_t = \varepsilon_t^*$, distributed $N(0, 1)$ independently of $\xi_{t-1}, \xi_{t-2}, \ldots$

When $a = 1$, a realization of this sequence is quite unlike a realization of Hypothesis A with the same value for ρ . But when a is close to ρ , realizations differ in appearance only subtly: with Hypothesis B occasional large jumps are more frequent than with Hypothesis A. Something like this kind of joint nonnormality is sometimes observed in practice.

An alternative hypothesis involving joint normality but incorrect specification is

Hypothesis C: $\{\xi_t\}$ is a jointly normal stationary autoregressive sequence of order greater than 1, or a jointly normal stationary moving-average sequence.

In each case, if ρ is the lag-1 serial correlation coefficient, let

$$\eta_t = \xi_t - \rho\xi_{t-1} .$$

Then as $n \to \infty$ the kurtosis statistics converge in probability:

$$b_{21} \to \frac{E(\eta_t^4)}{(E\eta_t^2)^2}, \quad b_{22} \to \frac{E(\eta_t^2 + \eta_{t+1}^2)^2}{(E\eta_t^2)^2}, \quad b_{23} \to \frac{E(\eta_t^2 + \eta_{t+1}^2 + \eta_{t+2}^2)^2}{(E\eta_t^2)^2}.$$

Under Hypothesis A these limits are 3, 8 and 15, respectively.

Under Hypothesis B,

$$\frac{E(\eta_t^4)}{(E\eta_t^2)^2} = 3 + \frac{12\rho^3(a-\rho)}{(1-\rho^2)^2},$$

$$\frac{E(\eta_t^2 + \eta_{t+1}^2)^2}{(E\eta_t^2)^2} = 8 + \frac{4\rho(a-\rho)(1+4\rho^2+a\rho^3)}{(1-\rho^2)^2},$$

$$\frac{E(\eta_t^2 + \eta_{t+1}^2 + \eta_{t+2}^2)^2}{(E\eta_t^2)^2} = 15 + \frac{4\rho(a-\rho)(2+a\rho+5\rho^2+a^2\rho^4)}{(1-\rho^2)^2}.$$

(The formidable polynomial manipulations have been computerized, and the above results are believed to be correct.)

For an asymptotic measure of power when n is large, the excesses of these expressions over the null-hypothesis values of 3, 8, 15 may be divided by the (conjectured) asymptotic standard deviations, $\sqrt{24/n}$, $\sqrt{112/n}$, $\sqrt{296/n}$, respectively. Suppose that $a > 0.9$ (say). Then b_{22} is more powerful than b_{21} when ρ is less than 0.75 about (the critical value varies a little with a); b_{22} is much more powerful when ρ is near 0; it is a little less powerful when ρ exceeds the critical value. And b_{23} is more powerful than b_{22} when ρ is less than 0.7 about.

Under Hypothesis C, η_t, η_{t+1}, η_{t+2} are jointly normally distributed with, in general, nonzero correlation. Let the lag-h serial correlation coefficient be δ_h . Then

$$\frac{E(\eta_t^4)}{(E\eta_t^2)^2} = 3, \quad \frac{E(\eta_t^2 + \eta_{t+1}^2)^2}{(E\eta_t^2)^2} = 8 + 4\delta_1^2,$$

$$\frac{E(\eta_t^2 + \eta_{t+1}^2 + \eta_{t+2}^2)^2}{(E\eta_t^2)^2} = 15 + 8\delta_1^2 + 4\delta_2^2.$$

The sampling distributions for the kurtosis statistics will be affected by the lack of independence of $\{\eta_t\}$, but the limiting value for b_{21} is not affected by the specification error. The limiting values for b_{22} and b_{23} are little affected if their excesses over the null-hypothesis values, namely $4\delta_1^2$ and $8\delta_1^2 + 4\delta_2^2$, are small compared with the respective standard deviations. The values of δ_1 and δ_2 can be estimated from the innovations $\{u_t\}$ as their lag-1 and lag-2 serial correlation coefficients.

References

F. J. ANSCOMBE (1961). Examination of residuals. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, **1**, 1-36.

F. J. ANSCOMBE and W. J. GLYNN (1975). Distribution of the kurtosis statistic b_2 for normal samples. Yale University, Department of Statistics, Technical Report No. 37.

R. A. FISHER (1930). The moments of the distribution for normal samples of measures of departure from normality. Proceedings of the Royal Society of London, A **130**, 16-28.

B. KLEINER, R. D. MARTIN and D. J. THOMSON (1979). Robust estimation of power spectra. Journal of the Royal Statistical Society, B, **41**, 313-351.

K. V. MARDIA (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, **57**, 519-530.

R. L. PLACKETT (1965). A class of bivariate distributions. Journal of the American Statistical Association, **60**, 516-522.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 47	2. GOVT ACCESSION NO. AD-A085 885	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TESTS FOR JOINT NORMALITY IN TIME SERIES		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Francis J. Anscombe and Ho-len H. Chang		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0563
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics, Yale University Box 2179, Yale Station New Haven, CT 06520		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-242
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research		12. REPORT DATE May, 1980
		13. NUMBER OF PAGES i + 11 + i = 13
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

To be presented at the annual statistical meetings, Houston, TX, August, 1980.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Kurtosis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Usual methods for analysis of time series, in time or frequency domains, are based on second-moment statistics, and are perhaps inappropriate if all variables are not jointly normally distributed. We propose an adaptation of Mardia's multivariate kurtosis test applicable to a sample from a stationary random sequence. An autoregressive structure of low order is estimated from the sample and used to estimate "innovations". A kurtosis measure may be defined from single innovations, or from pairs of consecutive innovations, or from triples of consecutive innovations, etc. Null-hypothesis distributions and power are investigated.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)